



Software for
Business Intelligence

BizInt Smart Charts

COOKBOOK OF REPORTS & VISUALIZATIONS

www.bizint.com/Cookbook

Created with the BizInt Smart Charts Product Family

APRIL 2019

using this COOKBOOK

APPLYING THESE RECIPES TO YOUR DATA

We believe the methods in the recipes included here are applicable to other data sets and sources. However, because each recipe was built using a particular example dataset, you should take some caveats into consideration when applying recipe steps to your own data:

Caveat One: When you combine charts from different sources, the source selected first (the “Key Chart”) will determine the names for common columns. For example, depending on the source, the column could be called Condition, Disease or Indications. While you may see different column names than those indicated in the recipe, the same steps should still apply.

Caveat Two: Some recipes contain steps and guidance based on the content of the example dataset. For example, we provide suggestions for changing Common Trial ID assignment in cases where the automatically generated Common ID combines trials we want to be treated separately. How much this step is necessary depends on the dataset.

Caveat Three: We revise these recipes based on user feedback, so check the latest edition of the Cookbook on our website to see if we have revised the recipe you're interested in.

STYLE CONVENTIONS USED

Commands are shown in **colored sans-serif type** with the pipe(|) symbol used to divide sequential menu selections. For example, **File | Combine** indicates that you should first click on the File menu and then select Combine.

Italics indicate a note, generally indicating steps that may be different depending on the dataset.

“Text in quotation marks” indicates something you should type in to a text box or text that appears in the software outside of the menus.

Note: as of December 2017, support for drug pipeline and clinical trials data is in the BizInt Smart Charts Drug Development Suite.

Note:

*This edition of the Cookbook is a selection of some of our favorite recipes, updated for **VantagePoint - Smart Charts Edition Version 11.***

If you are using an earlier version of VP-SCE, please refer to the February 2018 or earlier versions of the Cookbook, as many VP-SCE functions have been redesigned in Version 11.



THE JOURNEY CONTINUES...

TABLE OF CONTENTS

PIPELINE

Extracting the Highest Phase for an Indication [Pipeline DP-4].....	2
Highest Phase for an Indication – Database Summary [Pipeline DP-5]	4
Creating a Pipeline “Piano Chart” in PowerPoint [Pipeline DP-13].....	6
Visualizing the Pipeline with a Bullseye [Pipeline DP-16]	8

CLINICAL TRIALS

Training (1): Creating a Trials “Piano” Chart [Trials CT-11].....	10
Training (2): Trial Completion Timeline [Trials CT-12].....	12
Summarizing Clinical Trial Enrollment [Trials CT-4]	14

PATENTS

Analyzing Companies by Therapeutic Activity with a Bubble Chart [Patents PT-12].....	16
Summarizing Sequence Locations in a Patent Family [Patents PT-8]	18

TECHNIQUES

Technique: Using List Cleanup.....	20
Technique: Working with Thesauri.....	22
Technique: Extracting Keywords	24

About this Cookbook	26
---------------------------	----

Extracting the Highest Phase for an Indication [Pipeline DP-4]

In this example, we've used VantagePoint - Smart Charts Edition to normalize indications and phase terminology, and to create a copy of the Drug Development Phase subtable filtered for a specific indication (merkel cell carcinoma). In BizInt Smart Charts Drug Development Suite, we then run the new "Extract High Phase in Column" tool to extract the highest phase from the filtered subtable into a new column.

This revised recipe features the new "Filter Drug Development Phase" script, which makes filtering the Drug Development Phase table much simpler. The script also normalizes phase terminology to Arabic numbers.

SUBTABLE FILTERED FOR MCC

NORMALIZED HIGHEST PHASE

Combined: Merkel Cell Carcinoma 2017 - Citeline, Adis, Cortellis									
	Drug	Drug Development Phase			Drug Development Phase:Merkel cell carcinoma			Highest Phase (1)	Extracted Highest Phase
		Indication	Phase	Country	Indication (MCC)	Phase (MCC)	Country		
1	ALT 803	Bladder cancer	Phase I/III	USA	Merkel cell carcinoma	Phase 2	USA	Phase 2/3	Phase 2
		Merkel cell carcinoma	Phase II	USA					
		Acute myeloid leukaemia	Phase II	USA					
		Multiple myeloma	Phase I/II	USA					
		Non-small cell lung cancer	Phase I/II	USA					
		Haematological malignancies	Phase I/II	USA					
		Non-Hodgkin's lymphoma	Phase I/II	USA					
		Pancreatic cancer	Phase I/II	USA					
		Malignant melanoma	Phase I	USA					
		HIV infections	Phase I	USA					
Bacterial infections	Preclinical	USA							
2	Activated natural killer cells - NantKwest	Merkel cell carcinoma	Phase II	USA	Merkel cell carcinoma	Phase 2	USA	Phase 2	Phase 2
		Haematological malignancies	Phase I	Canada					
		Renal cancer	Phase I	USA					
		Solid tumours	Phase I	Germany					
		Acute myeloid leukaemia	Phase I	USA					
		Malignant melanoma	Phase I	USA					
	BGB-A317	Cancer, bladder	Phase II		Merkel cell carcinoma	Phase 1		Phase 2	Phase 1
		Cancer, lymphoma, Hodgkin's	Phase II						
		Cancer, solid, unspecified	Phase I						
		Cancer, renal	Phase I						
		Cancer, breast	Phase I						
		Cancer, brain	Phase I						
		Cancer, colorectal	Phase I						

HIGHEST PHASE COLUMN EXTRACTED FROM THE FILTERED SUBTABLE

Extracting the Highest Phase for an Indication [Pipeline DP-4]

How did we create this?

SEARCHES: Cortellis, Adis R&D Insight, and Citeline Pharmaprojects: Indication = Merkel Cell Carcinoma

BIZINT SMART CHARTS DRUG DEVELOPMENT SUITE - *create your combined chart*

1. Import each set of search results into BizInt Smart Charts (select any template) and use **File | Combine** features to create a combined chart. Use the Adis R&D Insight chart as the key chart so your drug names match the ones used in the recipe.
2. Use **Tools | Identify Common Drug Names** to create the Common Drug Name column and use **View | Sort** to sort by this field.
3. Use **View | Columns** to make the following fields visible: Drug, Highest Phase, and Drug Development Phase.
4. Use **File | Export** to export as VantagePoint - Smart Charts Edition. VantagePoint - BizInt Smart Charts Edition (VP-SCE) should open and import the file.

VANTAGEPOINT - SMART CHARTS EDITION - *filter the Drug Development Phase for indication of interest*

5. If the file does not launch automatically in VantagePoint - Smart Charts Edition (VP-SCE), then import the .vpxi file into VP-SCE using **Import from BizInt Smart Charts**, under the **Home** tab.
6. **Clean up your selected indication:** On the Summary panel, double-click on Indication under Drug Development Phase.
7. If there are Indications containing data in parentheses, under the **Refine** tab, select **Further Processing | Read Until - First | Paren** to remove qualifiers in parenthesis..
8. Click on **List Cleanup**. "Field to Clean" is the field created in the previous step (e.g. Drug Development Phase::Indication: Read Until First/Paren). For Matching ruleset, select General from the drop-down and click OK..
9. Enter "merkel" into either search box (overwrite "(type to filter...)" to filter the list. Drag and drop all the Merkel Cell Carcinoma items onto a single item.
For more information on using List Cleanup, see the Techniques: Using List Cleanup recipe.
10. To create the filtered subtable, you do not need to cleanup any other indications beyond Merkel Cell Carcinoma, so just click **Accept**. Saving the thesaurus is optional.
11. **Filter the subtable by your Indication and normalize phase terminology:** You will now have a new sub-field: Indication:: Read Until - First/Paren (Cleaned). Select the cell with your cleaned indication, "Merkel cell carcinoma" in this case.
12. Under the **Analyze** tab, click on **Run a Script | Filter Drug Development Phase**. A new filtered drug development phase field will be created along with a normalized phase field (e.g. Phase (1)).
13. **Rename the subtable fields:** Go to the Summary Sheet and right click on each of the subfields of "Drug Development Phase: Merkel cell carcinoma" and select **Rename Field...** Rename the indication field to "Indication: MCC" and the phase field to "Phase (MCC)". *Note: subtable fields cannot be changed in BizInt Smart Charts, so you need to rename them in VP-SCE.*
14. **Export filtered subtable and normalized phases back to BizInt Smart Charts:** Under the **Report** tab click on **Export to BizInt Smart Charts**. In the Export Wizard, uncheck the boxes for all fields except the filtered drug development phase (e.g. Drug Development Phase: Merkel cell carcinoma) and Highest Phase (1). Click **Export**.

BIZINT SMART CHARTS DRUG DEVELOPMENT SUITE- *extract the highest phase for the indication into a new column*

15. When BizInt Smart Charts Drug Development Suite comes to the foreground, select the combined chart . It will redisplay with two new columns -- the new **filtered Drug Development Phase** subtable and a new **Highest Phase (1)** column.
16. **Extract the highest phase for your Indication into a new column:** Select the **filtered Drug Development Phase** subtable and then select **Tools | Extract High Phase in Column**. A new column will be created showing the Extracted Highest Phase from the filtered MCC subtable.
You can compare this to the Highest Phase (All Indications) column to see where the highest phase for Merkel Cell Carcinoma differs from the highest phase for all indications.

Highest Phase for an Indication – Database Summary [Pipeline DP-5]

In this example, we've taken the chart from the previous recipe (Extracting the Highest Phase for an Indication) and used BizInt Smart Charts Reference Rows to select a single value for each unique drug. And, we've used tools in BizInt Smart Charts Reference Rows to create a summary of the Merkel Cell Carcinoma Phase and Update Date for each related drug record.

Combined: Merkel Cell Carcinoma 2017 - Citeline, Adis, Cortellis								
	Drug	Database	Highest Phase (All Indications)	Highest Phase (MCC)	Database - MCC Phase - Update Date			
					Database	Highest Phase (MCC)	Last Update	
1.	cabozantinib S-malate	1.1 CORTL link	Launched	Phase 2	Cortellis from Clarivate Analytics	Phase 2	2017-07-21	1.1
		1.2 Adis link			Adis R&D Insight	Phase 2	2017-07-31	1.2
		1.1 CORTL	1.2 Adis	1.2 Adis				
2.	pembrolizumab	2.1 CORTL link	Launched	Phase 2	Cortellis from Clarivate Analytics	Phase 2	2017-08-05	2.1
		2.2 Adis link			Adis R&D Insight	Phase 2	2017-08-04	2.2
		2.3 Pipeln link			Citeline Pharmaprojects	Preclinical	2017-08-04	2.3
		2.1 CORTL	2.1 CORTL	2.1 CORTL				
3.	sapanisertib	3.1 CORTL link	Phase 2	Phase 1/2	Cortellis from Clarivate Analytics	Phase 2	2017-06-21	3.1
		3.2 Adis link			Adis R&D Insight	Phase 1/2	2017-06-27	3.2
		3.1 CORTL	3.2 Adis	3.2 Adis				
4.	pasireotide pamoate (sustained release), Novartis	4.1 CORTL link	Launched	Phase 1	Cortellis from Clarivate Analytics	Phase 1	2017-07-25	4.1
		4.2 CORTL link			Cortellis from Clarivate Analytics	Phase 1	2017-07-25	4.2
		4.3 Adis link			Adis R&D Insight	Phase 1	2017-07-31	4.3
		4.4 Pipeln link			Citeline Pharmaprojects	Phase 1	2017-07-31	4.4
		4.1 CORTL	4.3 Adis	4.3 Adis				
5.	ALT 803	5.1 Adis link	Phase 2/3	Phase 2	Adis R&D Insight	Phase 2	2017-07-17	5.1
		5.1 Adis	5.1 Adis	5.1 Adis				

SUMMARY OF MCC PHASE BY DATABASE AND UPDATE DATE

Highest Phase for an Indication – Database Summary [Pipeline DP-5]

How did we create this?

SEARCHES: Cortellis, Adis R&D Insight, and Citeline Pharmaprojects: Indication = Merkel Cell Carcinoma

BIZINT SMART CHARTS DRUG DEVELOPMENT SUITE

1. Open your combined pipeline chart from the previous recipe [Pipeline DP-5] with the extracted highest phase for Merkel cell carcinoma column visible.
2. **Rename columns:** Right-click on the Extracted Highest Phase column and choose **Column Properties**. Rename the column “Highest Phase (MCC)” and click OK. Repeat to rename “Highest Phase (1)” to “Highest Phase (All Indications)”.
3. Save the .cht file.
4. Send the chart to Reference Rows using the toolbar icon (or **File | Send to Reference Rows**).

BIZINT SMART CHARTS REFERENCE ROWS

5. During the Create Reference Rows wizard, make sure that Cortellis is listed as the first database in Step 2. Click Next.
6. **Apply rules to select most recently updated records:** On the third panel, select the Highest Phase (All Indications) column, click the Edit button and choose the Most Recently Updated rule to select the phase from the most recently updated record. Do the same for the Highest Phase (MCC) column, and click Finish.
7. When the wizard finishes, you will see the Selection View. Green check marks indicate the items selected by the database ranking and rules.
8. **Create a Summary Subtable:** In the Selection View, run **Tools | Create Subtable from Columns**.
9. Select the Database, Highest Phase (MCC) and Last Update columns in that order and press Apply. A new column will be created containing a subtable with these three columns.
10. Double click on the new subtable column header to rename the column from Untitled Subtable to “Database - MCC Phase - Update Date”.
11. **Apply the Summarize rule to the subtable:** Use **Reference Rows | Column Rules...** to set the rule for the new subtable column to Summarize All Values. Blue hash marks will appear in the cells. *Note: you will not see the effect of the Summarize rule until you export the report.*
12. Hide any columns you do not want displayed in your final report.
13. **Create the HTML report:** Use **File | Export** to export the report to HTML and see the selected highest phases and the subtable column summarizing the Database/Hep C Phase/Update data. *You can use this report to review reported phase discrepancies between databases.*
14. You can now sort the Reference Rows report by the Highest Phase for MCC or the Highest Phase for all indications, since a phase has been selected for the group of records for each drug. These selected values for each drug can also be exported to VantagePoint - Smart Charts Edition to use in visualizations, as you will see in the Piano Chart and Bullseye recipes in this Cookbook.

Creating a Drug Pipeline “Piano Chart” in PowerPoint [Pipeline DP-13]

This recipe describes how to build a piano chart visualization of a drug pipeline, which shows one box for each drug, boxes color-coded by field values, and boxes labeled by multiple fields. In this example we have color coded the boxes by delivery route and included both drug name and originator company in the boxes.

Search results from several pipeline databases were exported from InfoDesk Pipeline Plus to BizInt Smart Charts Drug Development Suite and then integrated with Reference Rows. In Reference Rows we applied columns rules and reviewed the data to select representative values for each drug. VantagePoint - Smart Charts Edition (VP-SCE) was used to clean-up and simplify drug names, company names, delivery route, and phases.

We then use VP-SCE to automatically build the chart in PowerPoint (or Excel) using the “Piano Chart” dialog shown below. The chart can then be edited in PowerPoint for final presentation.

MCC Drugs by Phase & Delivery Route

Biological Testing	Preclinical	Clinical, Phase Unknown	Phase I	Phase II	Launched	Discontinued	No Development Reported
Merkel cell carcinoma therapies Vironika	OK-301 TG Therapeutics	avelumab companion diagnostic Merck KGaA/Pfizer	BOB-A317 Celgene	ALT 803 Altor BioScience Corporation	avelumab Merck KGaA	torvotuzumab mertansine ImmunoGen	ATN-161 Attenuon
	ETBX-051 NantWorks		ID-G100 Immune Design	CST-101 NantWorks			tivantinib ArQule
	ETBX-061 NantWorks		pasireotide Novartis	F16-IL2 Phlogien			
	ipilimumab Bristol-Myers Squibb Co		SIRPa-Fc Trillium Therapeutics	MCPyV vaccine Fred Hutchinson Cancer Research Center			
	LTvx APCure		utomilumab Pfizer	nivolumab Ono Pharmaceutical Co Ltd			
Merkel cell polyomavirus inhibitors Vironika			ID-G100 Immune Design	pembrolizumab Merck & Co	http://bizint-pipeline.citeline.com/CpViewAccess.aspx?B=69763&C=0&D=12		
			SIRPa-Fc Trillium Therapeutics	PEN-221 Tarveda Therapeutics			
				tavokinogene tetsaplasmid OncoSec Medical			
				talmogene laherparepvec BioVex inc			
				tavokinogene tetsaplasmid OncoSec Medical			
				cabozantinib S-malate Exelixis			
				pazopanib GlaxoSmithKline plc			
				sapanisertib Intellikine			

Piano Chart

Select a Template:
 Select a Template

Step 1: Select Required Fields:
 Select a Field for the Box Labels:
 Product

Select a Field for the Phase:
 Highest Phase (Indication) Show All Items

Step 2: Choose Options:
 Select a Field to add to the Boxes (optional):
 Company

Choose a Field to Color Code the Boxes:
 Route of Administration (Cleaned)

Add hyperlinks to drug records

Step 3: Set Chart Title:
 Chart Title: MCC Drugs by Phase & Delivery Route

Step 4: Create Piano Chart
 Choose the export format for your chart: PowerPoint Excel

Template Name: Show Template

Create Piano Chart Cancel

Route of Administration (Cleaned)

- Diagnostic
- Injectable
- Injectable, intratumoral
- Oral
- Unknown

Creating Pipeline “Piano Charts” in PowerPoint [Pipeline DP-13]

How did we create this?

This recipe assumes basic knowledge of using the BizInt Smart Charts Drug Development Suite and Reference Rows tools. For more information on using the features below, refer to the BizInt Smart Charts Drug Development Suite Mini Guide, or contact support@bizint.com.

BIZINT SMART CHARTS DRUG DEVELOPMENT SUITE - create your combined chart

1. Import each set of search results into BizInt Smart Charts and use **File | Combine** features to create a combined chart
2. Use **Tools | Identify Common Drug Names** to create the Common Drug Name column.
3. Use **View | Columns** to make the following fields visible: drug name, overall highest phase, company/originator, and Route of Delivery/Administration.
4. Save the chart, and select **File | Send to Reference Rows** to open the chart in BizInt Smart Charts Reference Rows.

BIZINT SMART CHARTS REFERENCE ROWS - build a single row per drug to send to VP-SCE

5. Click Next to skip regenerating Common Drug Names. Select your desired Database Ranking in the wizard, then click Finish to see the Selection View.
6. Apply Column Rules (for example, Most Recently Updated for the phase field, Least Content (characters) for the drug name and company/originator name).
7. **Choose content to include in the piano chart:** Review the selections (indicated by green checkmarks) and edit as needed, either by right-clicking on a different cell and choosing **Select this cell** or by double-clicking in a cell to edit the contents.
8. Use **File | Export** to export to VantagePoint - Smart Charts Edition. VantagePoint - BizInt Smart Charts Edition (VP-SCE) should open and import the file.

VANTAGEPOINT - BIZINT SMART CHARTS EDITION - clean fields and build the piano chart

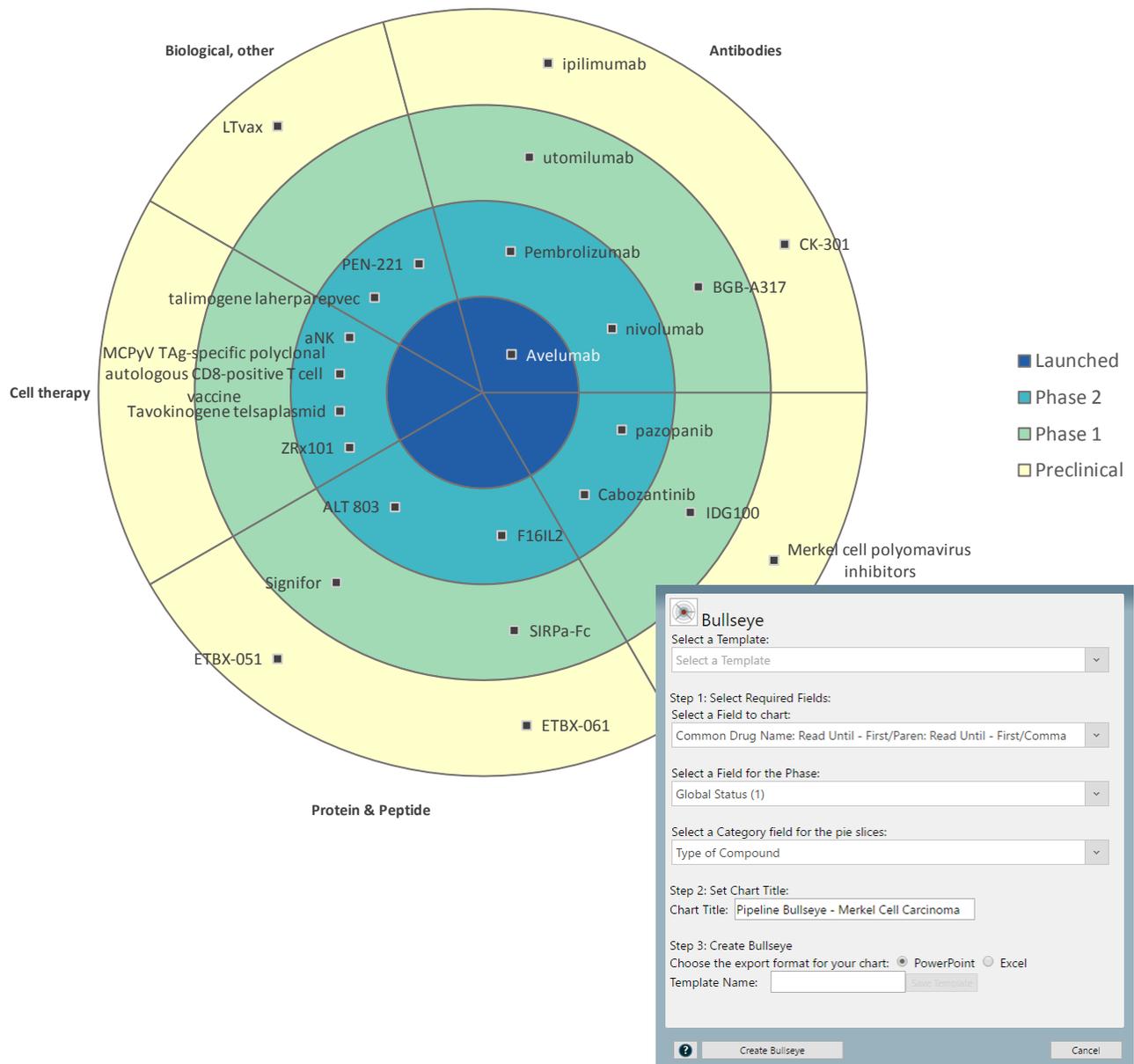
9. **Clean up fields used in the Piano Chart (delivery route and company):** On the Summary screen, right-click on the field name and select List Cleanup. For Matching ruleset, select General (for delivery route) or Organization Names (Pharma) (for company/originator names) from the drop-down and click OK.
Use the Cleanup Confirmation panel to clean-up terms -- drag and drop to combine and then click Accept. *Saving the thesaurus is optional.*
Note that a box will be created for each drug and originator pair. If you have more than one originator for a drug you will see multiple boxes for that drug. If you have more than one delivery route per drug, only the first will be used to color the boxes.
10. **Clean up (shorten) drug names:** Go to Home | List and choose your drug name field and click OK. If there are names with data in parentheses, under the **Refine** tab, select **Further Processing | Read Until - First | Paren** to remove indication from drug names . Then use **Further Processing | Read Until - First | Comma** to remove company names from drug names.
11. **Set Phase as the Data Type for your phase field:** In the Summary Sheet, right click on the phase field, and from the “Set Data Type” submenu, select “Custom” and then “Phase”. *Note: This tells VP-SCE to sort phases correctly in your visualization.*
12. **Clean up phases:** Right-click on the highest phase field and select Thesaurus. Under Thesaurus drop-down, select DrugDevelopmentPhases_Arabic and click OK.
13. **Build the Piano Chart:** Go to the **Report** tab and click on **Piano Chart**.
In Step 1, select the fields you want to include in your chart. Choose the cleaned drug name field for the Box Labels and your cleaned phase field for the Field for the Phase.
14. In Step 2, choose the cleaned originator field as the “Field to add to the Boxes” and the cleaned delivery route as the “Field to Color Code the Boxes”. Select the check box to add hyperlinks to the boxes if desired.
Note: record links can be followed in PowerPoint in the Slide Show view.
15. In Step 3, enter a title for your chart. In Step 4, select the radio button for Excel or PowerPoint and click “Create Piano Chart”.
16. VP-SCE will create a new document in PowerPoint or Excel and draw the piano chart. Go to PowerPoint or Excel to view the chart.
You can edit the chart in PowerPoint. In Excel, use the option to Copy as Picture to create an image that you can paste into another application. You can also save the selection as a PDF from Excel.

Visualizing the Pipeline with a Bullseye

[Pipeline DP-16]

This recipe shows how to create our newest visualization -- a bullseye chart created in PowerPoint.

The bullseye shown here was created by integrating Merkel Cell Carcinoma pipeline data from Cortellis and Pharmaprojects. BizInt Smart Charts and VantagePoint tools were used clean up drug names and phases, and to create a "type of compound" category for the sectors or slices of the chart. Some review by the analyst is also required. Once the data is prepared, the bullseye is drawn automatically by VantagePoint - Smart Charts Edition as a PowerPoint slide, which can then be edited if needed for final presentation.



Bullseye

Select a Template:

Step 1: Select Required Fields:

Select a Field to chart:

Select a Field for the Phase:

Select a Category field for the pie slices:

Step 2: Set Chart Title:

Chart Title:

Step 3: Create Bullseye

Choose the export format for your chart: PowerPoint Excel

Template Name:

Visualizing the Pipeline with a Bullseye [Pipeline DP-16]

How did we create this?

The **type of compound.the** file used in this recipe can be found at: www.bizint.com/vptraining

SEARCHES: Cortellis and Pharmaprojects: Indication = Merkel Cell Carcinoma

BIZINT SMART CHARTS DRUG DEVELOPMENT SUITE

1. **Import data:** Use **File | Import** to import the sample data from each database. Choose the General Product Chart (Pharmaprojects) or Basic Information (Cortellis) as the initial chart template.
2. **Combine data:** Use **File | Combine** to combine both charts, choosing the Citeline chart as your key chart.
3. **Display key information:** Select **View | Columns** to add the columns you plan to use – for this recipe, include Global Status, Origin of Material, and Technologies columns.
4. Save the chart, and select **File | Send to Reference Rows** to open the chart in BizInt Smart Charts Reference Rows.

BIZINT SMART CHARTS REFERENCE ROWS

5. In the wizard, click Generate (to generate Common Drug Names.) Then click Next and on the next step, click Finish.
6. Use **File | Export** to export as VantagePoint - Smart Charts Edition. VantagePoint - BizInt Smart Charts Edition (VP-SCE) should open and import the file.

VANTAGEPOINT - SMART CHARTS EDITION - *clean up the data*

7. **Create a field combining different publisher fields:** Under the **Refine** tab, click on **Merge Fields**. For this recipe, select and add the Origin of Material and Technologies fields and click OK.
8. **Apply thesaurus to create categories:** Click on **Thesaurus** and enter "Type of compound" for the Name of new field. Select the "Type of Compound" thesaurus. Click on More Options and uncheck the box to "Include unmatched items in new field." Click OK.
9. **Export back to Reference Rows for review:** Under the Report tab click on Export to BizInt Smart Charts. In the Export Wizard, select only the "Type of Compound" field and Click Export.

BIZINT SMART CHARTS REFERENCE ROWS - *review and select data*

10. When BizInt Smart Charts Reference Rows comes to the foreground, select the combined chart. It will redisplay with the new column "Type of Compound." Review and edit the column so that there is only one Type of Compound value for each record.
11. Use **File | Export** to export to VantagePoint - Smart Charts Edition.

VANTAGEPOINT - SMART CHARTS EDITION - *clean up the data and create the Bullseye chart*

12. **Set Phase as Data Type:** If not already set, right-click on the Global Status field and choose **Set Data Type | Custom | Phase**.
13. **Normalize Phase:** Right-click on Global Status field and choose **Thesaurus**. For the Thesaurus, select "DrugDevelopmentPhases_Arabic" and click OK.
14. **Create sub-dataset:** Double-click on the column header to sort by phase. Use Shift+click to select Preclinical through Launched. Click on **Create Sub-dataset** and click OK. *A new dataset appears with your selected phases.*
15. **Clean Common Drug Name:** Right-click on Common Drug Name and use **Further Processing | Read Until - First | Paren** to remove indication from drug names.
16. **Create Bullseye:** Go to the Report tab and click on **Bullseye**.
17. In Step 1, select the cleaned Common Drug Name as the Field to chart. Select Global Status(1) as the Field for the Phase and Type of compound as the Category field for the pie slices. Type your chart title in Step 2.
18. In Step 3, select PowerPoint as the export format and click **Create Bullseye**. Some intermediate windows will appear and then VP-SCE will create a new document in PowerPoint and draw the bullseye, similar to the one shown on the previous page.
Note: You can edit the chart in PowerPoint and follow links to the trial records in the PowerPoint show.

Training (1): Creating a Trials “Piano” Chart

[Trials CT-11]

This is the first in a set of two training recipes designed to help BizInt Smart Charts users learn the key features of VantagePoint - Smart Charts Edition (VP-SCE). In these two recipes, you will learn how to clean-up data and create two visualizations — a Trials Piano Chart and an Estimated Trial Completion Timeline.

These recipes use a **training.vpxi** file which was created by searching for Merkel Cell Carcinoma trials in ClinicalTrials.gov and EU Clinical Trials (EudraCT), importing the results into the BizInt Smart Charts Drug Development Suite and then integrating results for each trial in BizInt Smart Charts Reference Rows. In Reference Rows we applied column rules and reviewed the data to select representative values for each trial and then exported to create the training.vpxi file.

MCC Trials

Phase I	Phase I/II	Phase II
NCT00655655 National Cancer Institute (NCI)	NCT01440816 EMD Serono	2011-003226-27 GlaxoSmithKline
NCT02831179 National Cancer Institute (NCI)	NCT02488759 Bristol-Myers Squibb	2014-004735-39 Novartis
NCT02890368 Trillium Therapeutics Inc.	NCT02514824 Millennium Pharmaceuticals, Inc.	2016-000461-23 Bristol-Myers Squibb
NCT03000257 AbbVie	NCT02643303 MedImmune LLC	NCT01013779 Trans-Tasman Radiation Oncology Group (TROG)
NCT03071757 AbbVie	NCT02936323 Tarveda Therapeutics	NCT02036476 Exelixis
NCT03107663 ImaginAb, Inc.	NCT03126110 Incyte Corporation	NCT02155647 Merck KGaA
NCT03212404 Checkpoint Therapeutics, Inc.	NCT03167164 NantCell, Inc.	NCT02196961 Bristol-Myers Squibb
	NCT03241173 Incyte Corporation	NCT02267603 National Cancer Institute (NCI)
		NCT02351128 University Hospital, Grenoble
		NCT02465957 NantKwest, Inc.
		NCT02819843 Amgen
		NCT02978625 National Cancer Institute (NCI)
		NCT03071406 Bristol-Myers Squibb

Piano Chart

Select a Template:

Step 1: Select Required Fields:
 Select a Field for the Box Labels:

Select a Field for the Phase:

Step 2: Choose Options:
 Select a Field to add to the Boxes (optional):

Choose a Field to Color Code the Boxes:

Add hyperlinks to drug records

Step 3: Set Chart Title:
 Chart Title:

Step 4: Create Piano Chart
 Choose the export format for your chart: PowerPoint Excel

Template Name: Save Template

Training (1): Creating a Trials “Piano” Chart [Trials CT-11]

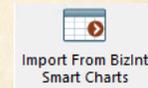
How did we create this?

This recipe is designed to be used with a *training.vpxi* file and provides detailed step-by-step instructions for cleaning up the data and creating a piano chart as shown on the previous page.

The **training.vpxi** file used in this recipe can be found at:
www.bizint.com/vptraining

VANTAGEPOINT - BIZINT SMART CHARTS EDITION - *Clean-up the data*

- Import data:** Open VantagePoint - Smart Charts Edition (VP-SCE) and click **Import from BizInt Smart Charts**. Choose the **training.vpxi** file and click Open.
- Clean-up sponsor names:** On the Summary Sheet screen, right-click on the Sponsor field and select **List Cleanup**. Under matching ruleset, select Organization Names (Pharma) and click OK.
- Use the Cleanup Confirmation panel to clean-up terms. On either panel, drag and drop names to combine terms (for example, drag “Incyte Biosciences International...” onto “Incyte Corporation” to combine these terms). When you are done with clean-up, click the Accept button. *You do not need to save the Thesaurus. A new field “Sponsor (Cleaned)” is created.*
For more information on using List Cleanup, see the Techniques: Using List Cleanup recipe.
- Return to the Summary Sheet. *Note: The Summary Sheet is where you can navigate between lists. You can right-click on a field name and use Rename Field to rename a field.*
- Set Phase as the Data Type for your phase field:** Right click on the phase field, and from the “Set Data Type” submenu, select “Custom” and then “Phase”. *This will cause phases to sort properly in your visualization.*
- Normalize phase:** Right-click on the phase field and choose **Thesaurus**. Under Thesaurus, select DrugDevelopmentPhases_Arabic and click OK. *Note: if you don’t see this file, click on the folder icon and go up one level to VantagePoint - SCE and select the Thesaurus folder.*
- Normalize status:** Return to the Summary Sheet. Right-click on the Overall Status field and choose Thesaurus. Select the “TrialStatus.the” file and click OK. *You will now have two more cleaned up fields “Phase (1)” and “Overall Status (1)”.*
- Having cleaned up sponsors, phases and trial status, you are ready to build your visualization.



VANTAGEPOINT - BIZINT SMART CHARTS EDITION - *Create the visualization*

- Build the Piano Chart:** Go to the **Report** tab and click on **Piano Chart**.
- In Step 1, choose the **Common Trial ID** field for the Box Labels and choose your cleaned-up **Phase (1)** field for the Phase.
- In Step 2, choose the cleaned Sponsor (Cleaned) field to add to the Boxes and the normalized Overall Status (1) field as the Field to Color Code the Boxes. Select the check box to add hyperlinks to trial records.
Note: record links can be followed in PowerPoint in the Slide Show view.
- In Step 3, enter the title “MCC Trials” for your chart. In Step 4, select the radio button for PowerPoint and click “Create Piano Chart”. *Your Create Piano Chart window should look as shown on the previous page.*
- Some intermediate windows will appear and then VP-SCE will create a new document in PowerPoint and draw the piano chart, similar to the one shown on the previous page. *Switch to PowerPoint to see the Piano Chart.*
Note: You can edit the chart in PowerPoint and follow links to the trial records in the PowerPoint show.
- To save your work in VP-SCE, go to the **Home** tab and select **Save** to save as a .vpt file.

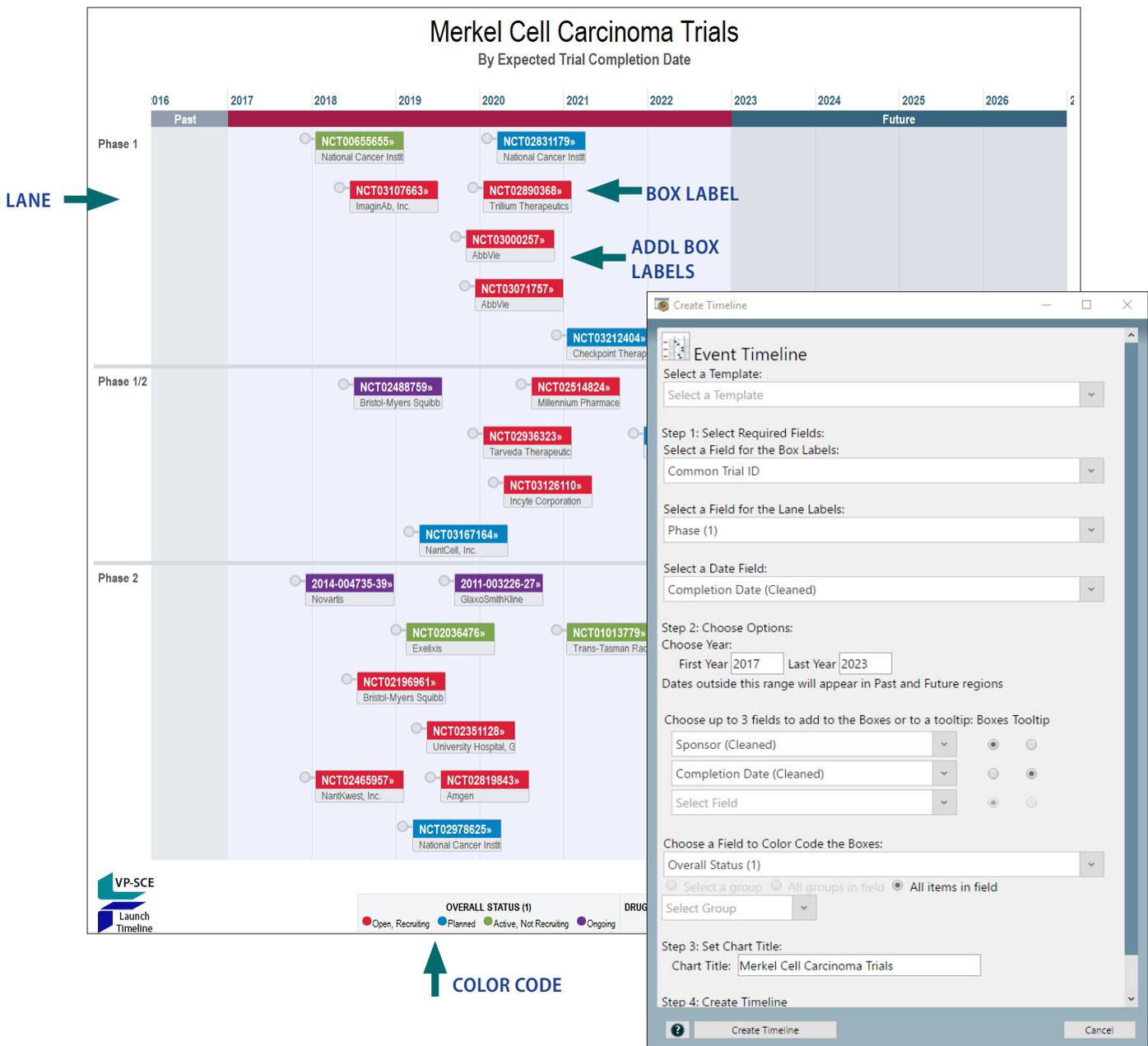
Note: If you have problems at any step, go back and check that you did not skip a step.

Do not hesitate to contact us at “support@bizint.com” for help!

Training (2): Trial Completion Timeline

[Trials CT-12]

In this second recipe, you will clean-up trial completion dates and use your previously cleaned data to create a trial completion timeline. The timeline shows estimated completion dates for a set of trials arranged into horizontal categories, or lanes. This type of chart is also sometimes called a “swimlane” or “fishbone” chart. The trial completion timeline is created in VP-SCE and can be saved as an interactive HTML report.



Training (2): Trial Completion Timeline [Trials CT-12]

How did we create this?

This recipe is a continuation of the first recipe in this series and uses the VP-SCE file created in the Training (1) recipe.

This recipe requires files which can be found at:

www.bizint.com/vptraining

FILES: FindReplace_MonthYYYY.the (Contact support@bizint.com for the files and installation directions)

VANTAGEPOINT - SMART CHARTS EDITION - *Clean-up completion dates*

1. **Remove text from completion date:** On the Summary Sheet, right-click on the completion date field and select **Further Processing | Read Until - First | Paren**.
2. **Convert partial dates:** Go to the Refine tab under Thesaurus select Find and replace. For the Name of the new field use "Completion Date: Read Until - First Paren". Under Find and Replace, select FindReplace_MonthYYYY from the drop-down list and then click OK.
3. **Rename the cleaned/converted completion date field:** Go to the Summary Sheet, right click on the cleaned field (probably "Completion Date: Read Until - First Paren (1)"), select **Rename Field...**, and rename the field "Completion Date (Cleaned)".
Note: this step is not required, but will make it easier for you to identify this field for the visualizaton.

VANTAGEPOINT - SMART CHARTS EDITION - *Create the visualization*

4. **Create the Timeline:** Go to the **Report** tab and click on **Event Timeline**. Then, select the fields that you want to include in your timeline.
 - Field for **Box Labels:** used for the box label - white text on colored background (e.g. Common Trial ID in this example).
 - Field for **Lane Labels:** used to label the lanes on the left of the timeline (e.g. Phase (1)).
 - Field for **Date:** the cleaned completion date, used to plot the boxes (e.g. Completion Date (Cleaned)).
 - First Year/Last Year: defines the date range of the launch timeline (e.g. 2017 and 2023).
 - Fields for boxes/tooltips: optional additional labels for the drug boxes (up to 3). These fields can also be used as "tooltips" when your mouse hovers over the drug box in the HTML version. (e.g Sponsor (Cleaned) with Boxes selected, and Completion Date (Cleaned) with Tooltip selected).
 - Field for color code: used to color code the trial boxes (e.g. Overall Status (1)). For the radio button, select "All items in field."
 - Type "Merkel Cell Carcinoma Trials" for the Chart Title. Your script panel should look like the previous page.
5. Click on the **Create Timeline** button to open your trial completion timeline in a VP-SCE window.
Note: The timeline adjusts to the size of the window so your timeline could look different than the timeline on the previous page.
6. You can edit the title, subtitle and other information in the HTML window or the VP-SCE window. In VP-SCE, you can click on each box to get more detail.
7. **Export the timeline:** To save the timeline, click the **Export Image** camera icon on the right side of the toolbar. You can save as HTML or as a PDF, JPG or PNG image.
Note: Timelines saved as HTML can be opened in a browser window and have some interactive features- you can zoom, scroll, see tooltips, and click on a trial id to view the record on CT.gov or EudraCT.
8. **Optional:** Try running the Create Timeline script again, but this time use Overall Status (1) for the Lane Labels and Phase (1) for the color code field, to create a slightly different visualization of the same data.



We hope these recipes have helped you learn some of the basic features of VantagePoint - Smart Charts Edition (VP-SCE). The next step is to try these recipes with your own data.

Summarizing Clinical Trial Enrollment

[Trials CT-4]

This recipe shows how you can use Reference Rows to build a single row for each trial that includes a summary table showing enrollment number and update date from each database. The chart used here is the same set of trials for drugs targeting PARP used in the trial timeline recipes in this Cookbook.

You can also create a summary table showing the enrollment in each member state for each trial. The EudraCT (EU Clinical Trials) database provides trial records for a trial in each member country. BizInt Smart Charts tools help you identify related trials and summarize the data for all EU trials in a single row.

Combined: CT.gov + EU trials: Merkel Cell Carcinoma

Trial Title	Database	Overall Status	Enrollment Summary				
			Database	Enrollment	Member State	Enrollment in State	
1. A Phase 1/2a, Open-label Multicenter Study to Assess the Safety, Tolerability, Pharmacokinetics, and Preliminary Antitumor Activity of PEN-221 in Patients With Somatostatin Receptor 2 Expressing Advanced Cancers	1.1 NCT link	Recruiting	ClinicalTrials.Gov	120 (Anticipated)		1.1	
	1.2 EUDRACT link		EUDRA Clinical Trials	120	UK - MHRA	40	1.2
	1.1 NCT	1.1 NCT					
2. Non-Comparative, Two-Cohort, Single-Arm, Open-Label, Phase 1/2 Study of Nivolumab (BMS-936558) in Subjects with Virus-Positive and Virus-Negative Solid Tumors	2.1 EUDRACT link	Ongoing	EUDRA Clinical Trials	220	Germany - PEI	27	2.1
	2.2 EUDRACT link		EUDRA Clinical Trials	220	Spain - AEMPS	38	2.2
	2.3 EUDRACT link		EUDRA Clinical Trials	220	UK - MHRA	22	2.3
	2.4 EUDRACT link		EUDRA Clinical Trials	220	Belgium - FPS Health-DGM	14	2.4
	2.5 EUDRACT link		EUDRA Clinical Trials	220	Netherlands - Competent Authority	8	2.5
2.1 EUDRACT	2.1 EUDRACT						
3. A Phase II Trial of STI-571/Imatinib (Gleevec®) (NSC-716051) in Neuroendocrine Carcinoma of the Skin (Merkel Cell Carcinoma)	3.1 NCT link	Completed	ClinicalTrials.Gov	40 (Actual)		3.1	
	3.2 NCT link		ClinicalTrials.Gov	80 (Actual)		3.2	
3.1 NCT	3.1 NCT						
4. An open label, multicenter pasireotide roll over protocol for patients who have completed a previous Novartis sponsored pasireotide study and are judged by the investigator to benefit from continued pasireotide treatment	4.1 EUDRACT link	Ongoing	EUDRA Clinical Trials	100	Greece - EOF	2	4.1
	4.2 EUDRACT link		EUDRA Clinical Trials	100	France - ANSM	9	4.2
	4.3 EUDRACT link		EUDRA Clinical Trials	100	Poland - Office for Medicinal Products	25	4.3
	4.4 EUDRACT link		EUDRA Clinical Trials	100	Spain - AEMPS	2	4.4
	4.5 EUDRACT link		EUDRA Clinical Trials	220	Portugal - INFARMED	4	4.5
	4.6 EUDRACT link		EUDRA Clinical Trials	100	Italy - Italian Medicines Agency	11	4.6
	4.7 EUDRACT link		EUDRA Clinical Trials	100	Belgium - FPS Health-DGM	22	4.7

Summarizing Clinical Trial Enrollment [Trials CT-4]

How did we create this?

SEARCHES: We searched EudraCT (EU Clinical Trials) and ClinicalTrials.gov for Merkel Cell Carcinoma. This is the same data used in the Training recipes to create piano charts and timelines.

BIZINT SMART CHARTS DRUG DEVELOPMENT SUITE

1. Import each set of search results into BizInt Smart Charts and use **File | Combine** to create a combined chart, selecting the ClinicalTrials.gov chart as the key chart (so that field names match the ones used later in the recipe.)
2. Run the **Identify Common Trial IDs** tool and sort by Common Trial ID. Review and make changes if needed.
3. Save the chart and send to Reference Rows using the toolbar icon (or **File | Send to Reference Rows**).

BIZINT SMART CHARTS REFERENCE ROWS

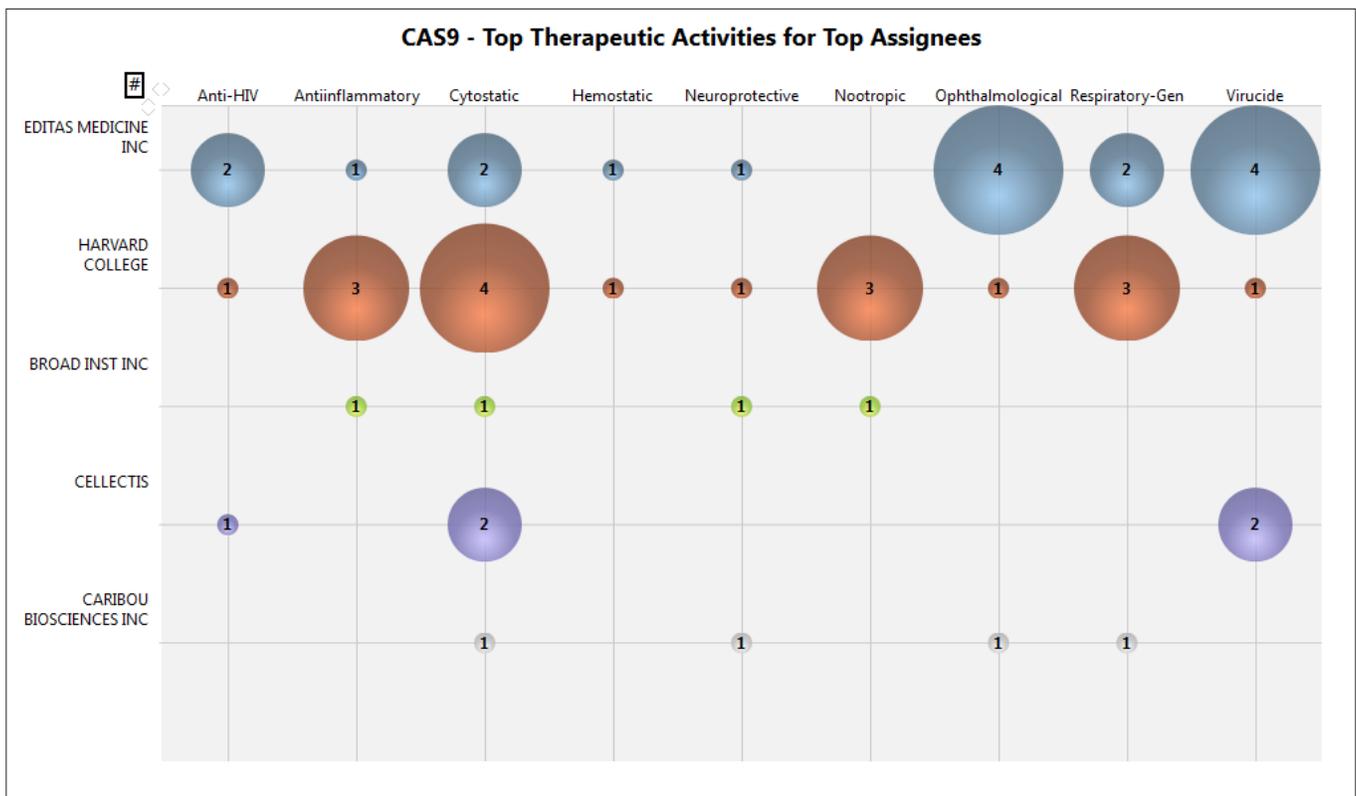
4. During the Create Reference Rows wizard, make sure that ClinicalTrials.gov is listed as the first database in Step 2, and then click Finish. You do not need to set any rules.
5. When the wizard finishes, you will see the Selection View. Green check marks indicate the items selected by the database ranking and rules.

BIZINT SMART CHARTS REFERENCE ROWS - *Create the Trial Enrollment Summary*

6. **Create the Enrollment Subtable:** In the Selection View, run **Tools | Create Subtable from Columns**.
7. Select the Database, Enrollment, Member State, and Enrollment in State columns in that order and press Apply. A new column will be created containing a subtable with these four columns.
8. Double click on the column header to rename the new column from Untitled Subtable to "Enrollment Summary".
9. **Change the column rule to summarize enrollment for each trial:** Use **Reference Rows | Column Rules...** to set the rule for the new column to **Summarize All Values**. Blue hash marks will appear in the cells.
10. Hide any columns you do not want displayed in your final report.
11. Use **File | Export** to export the report to HTML and see the summary subtable — see example at left. Both EU Clinical Trials and ClinicalTrials.gov provide total enrollment numbers; EudraCT adds the ability to see the enrollment in each EU country.

Analyzing Companies by Therapeutic Activity with a Bubble Chart [Patents PT-12]

This is one example of a bubble chart created with VantagePoint – Smart Charts Edition (VP-SCE) using data exported from BizInt Smart Charts for Patents (see other recipes in this Cookbook for bubble charts with pipeline and clinical trial data.) Here the bubble chart shows companies by therapeutic activity based on records from Derwent WPI, and PatBase. A single record for each patent family was created with BizInt Smart Charts Reference Rows and then exported to VP-SCE. We linked Probable Assignee from PatBase with Therapeutic Activity from DWPI to build the bubble chart in VP-SCE.



Analyzing Companies by Therapeutic Activity with a Bubble Chart [Patents PT-12]

How did we create this?

SEARCHES:

- GenomeQuest: Identify sequences in GenomeQuest for CRISPR/CAS9.
- PatBase & Derwent World Patents Index: Search using the patents numbers from the GenomeQuest sequence results.

BIZINT SMART CHARTS FOR PATENTS

1. Import each set of search results into BizInt Smart Charts and use the [File | Combine](#) feature to create a combined chart.
2. Run the [Identify Common Patent Families](#) tool and sort by the Common Family field.
3. Display the Probable Assignee and Therapeutic Activity columns.
4. Save the chart. Select [File | Send to Reference Rows](#) to open the chart in BizInt Smart Charts Reference Rows.

BIZINT SMART CHARTS REFERENCE ROWS

5. Use [File | Export](#) to export the visible columns to VantagePoint - Smart Charts Edition.

VANTAGEPOINT - SMART CHARTS EDITION

6. If the file does not launch automatically in VP-SCE, then import the .vpxi file into VP-SCE using [Import from BizInt Smart Charts](#) under the [Home](#) tab. Save your .vpt file.
7. **Cleanup Assignees:** Right-click on Probable Assignee and select List Cleanup... On the List Cleanup panel, under Matching ruleset, select Organization Names (Pharma) and click OK.
8. Use the Cleanup Confirmation panel to further cleanup company names and then click Accept. *See "Technique: Using List Cleanup" recipe for more details.*
9. Select the top assignees, then right click to select Add Selection to Group.... In the Add Items panel, type the word "top" and click OK. *To focus on industry, you can uncheck individuals and academic institutions.*
10. **Cleanup Therapeutic Activity:** Return to the Summary Sheet. Double-click on Therapeutic Activity to see values for this field. Under the [Refine](#) tab, select [Further Processing | Divide at | Semicolon](#). Then apply [Further Processing | Read Until - First | Period](#).
11. Select the top therapeutic activities, then right click to select Add Selection to Group.... In the Add Items panel, type the word "top" and click OK.
12. **Create the Bubble Chart:** Save the file. Under the [Report](#) tab, select [Bubble Chart](#) and a window will open with options for the Bubble Chart.
13. For "Step 1: Select Field for Rows", select the cleaned assignees field - Probable Assignee (Cleaned). Choose the "Select a group" radio button and then select the "top" group.
14. For "Step 2: Select Field for Columns", select your cleaned therapeutic activity field - Therapeutic Activity: Divide at/semicolon: Read Until - First/Period from the drop-down. Choose the "Select a group" radio button and then select the "top" group.
15. For Step 3, you can enter a template name if you want to save these settings.
16. Click "Run" and the a bubble chart showing assignees by therapeutic activity will be created.
17. **Edit the Bubble Chart:** You can click the "#" at the top left of the bubble chart to display the number of records in each bubble. Click on the chart title to edit it.
18. In VP-SCE, you can click on a bubble and see information about the records in each bubble in detail windows.
19. **Export the Bubble Chart:** Under the Report tab, click on Export Image to export the bubble chart in a variety of formats.

Summarizing Sequence Locations in a Patent Family [Patents PT-8]

In this example, we've used two new features in BizInt Smart Charts Reference Rows -- the "Create Subtable from Columns" tool and "Summarize All Values" rule -- to create a subtable summarizing key columns for sequences grouped by the Common Patent Family tool.

SUBTABLE SUMMARIZING RELATED SEQUENCES



Title	Database	Patent Assignee	Sequence Summary				
			Sequence ID	% Identity	Length	Location	
1. New diagnostic and therapeutic polynucleotides and polypeptides, useful in diagnosing a condition, disease or disorder associated with human molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or in gene mapping.	1.1 GENESEQ link	INCYTE CORP.	WO20040023973-5059	91.59	642	Claim 27; Page; 190pp; English.	1.1
	1.2 GPATPRT link		WO2004023973-5059	91.59	642	claim: 27; 28	1.2
	1.1 GENESEQ		1.1 GENESEQ				
2. New isolated PRO nucleic acid, useful for diagnosing, treating, or preventing, e.g. systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, systemic sclerosis, psoriasis, allergic disease, or asthma.	2.1 GENESEQ link	GENENTECH INC.	WO20050019258-1490	100.00	642	Claim 8; SEQ ID NO 1490; 158pp; English.	2.1
	2.2 GENESEQ link		WO20050016962-1490	100.00	642	Claim 8; SEQ ID NO 1490; 158pp; English.	2.2
	2.3 GPATPRT link		EP2014675-1490	100.00	642	probable disclosure (not found by automated parsing)	2.3
	2.4 GPATPRT link		US20070184444-1490	100.00	642	claim: 1; 2; 8; 19; 20; 21; 22; 23; 25; 26; 27	2.4
	2.5 GPATPRT link		US20110245090-1490	100.00	642	claim: 20; 23; 28; 31; 36; 38; 40	2.5
	2.6 GPATPRT link		WO2005016962-1490	100.00	642	claim: 2; 8; 19; 20; 21; 22; 23; 25; 26	2.6
	2.7 GPATPRT link		WO2005019258-1490	100.00	642	claim: 1; 2; 7; 19; 20; 21; 22; 23; 25; 26; 27	2.7
	2.8 GPATPRT link		EP2182006-1490	100.00	642	claim: 1; 3; 4; 5; 6; 7; 9; 10; 11	2.8
2.1 GENESEQ		2.1 GENESEQ					
3. Improved yeast cellular assay system for detecting e.g. polycyclic aromatic hydrocarbons, dioxins, PCBs comprising additional cellular proteins from the AHR signaling pathway which improve the sensitivity and maximal	3.1 GENESEQ link	WISCONSIN ALUMNI RES FOUND.		97.04	642	Example; Fig 6C; 26pp; English.	3.1
	3.2 GENESEQ link			99.84	642	Example; Fig 6C; 26pp; English.	3.2

Summarizing Sequence Locations in a Patent Family [Patents PT-8]

How did we create this?

SEARCHES:

- GenomeQuest – sequence search from GQPAT and/or GENESEQ.

BIZINT SMART CHARTS FOR PATENTS

1. Import each set of results into BizInt Smart Charts for Patents to create .chp files.
2. Use the **File | Combine** feature to create a single combined chart. Run the Identify Common Patent Families tool.
3. Display the Seq. ID Number, Percentage Identity, Subject Length, and Patent Sequence Location columns (**View | Columns**).
4. Change the title on the Patent Sequence Location column to “Location” (by double-clicking on the column header). Repeat to rename the Seq. ID Number column to “Sequence ID”, Percentage Identity to “% Identity”, and Subject Length to “Length”. *You need to give columns short names now, because you can't change them after creating the subtable in step 8, below.*
5. Save the .chp file.
6. Send the chart to Reference Rows using the toolbar icon (or **File | Send to Reference Rows**).

BIZINT SMART CHARTS REFERENCE ROWS

7. During the Create Reference Rows wizard, make GENESEQ the top-ranked database (if present). *This will cause the Derwent enhanced titles and normalized assignee to be preferred in the final report. This is faster than changing the database ranking in the columns one at a time*
8. When the wizard finishes, run **Tools | Create Subtable from Columns**.
9. Select the Sequence ID, % Identity, Length, and Location columns in that order and press Apply. A new column will be created.
10. Double click on the column header to rename the new column from Untitled Subtable to “Sequence Summary”.
11. Use **View | Columns** to hide the original Sequence ID, % Identity, Length, and Location columns.
12. Use **Reference Rows | Column Rules...** to set the rule for the new column to Summarize All Values.
13. Use **File | Export** to export the report to HTML and see the subtable column summarizing the sequence data.

Technique: Using List Cleanup

VP-SCE offers multiple ways to clean up data. The primary tool is List Cleanup found under the Refine tab. The majority of the recipes in this Cookbook include a List Cleanup step. Here we show some options for cleaning company names. The primary difference when cleaning up different fields is your choice of matching ruleset.

The screenshot displays the 'List Cleanup' tool interface. The 'Field to clean' is set to 'Originator: Read Until - First/Paren'. The 'Name of new field' is 'Originator: Read Until - First/Paren (Cleaned 2)'. The 'Matching ruleset' is 'Organization Names (Pharma)'. A list of rulesets is shown, including 'BritishAmericanSpelling', 'BritishAmericanSpellingStem', 'General', 'Organization Names (ignore dept)', 'Organization Names (Pharma)', 'Organization Names', 'Person Names-Aggressive', 'Person Names-Very Aggressive', and 'Person Names'. The 'Organization Names (Pharma)' ruleset is selected.

The 'Dual View Cleanup Confirmation' window shows a progress bar and a legend:

- Items in 1 record (14)
- Items in >1 Record (2)
- Sets (3)
- Cleaned up (3)

 The window displays two views of the data:

# Records	Item Name (16 items)
4	National Cancer Institute
2	AbbVie
1	Amgen
1	Checkpoint Therapeutics, Inc.
1	Exelixis
1	GlaxoSmithKline
1	ImaginAb, Inc.
1	MedImmune LLC
1	Millennium Pharmaceuticals, Inc.
1	NantCell, Inc.
1	NantKwest, Inc.
1	Novartis
1	Tarveda Therapeutics
1	Trans-Tasman Radiation Oncology Group
1	Trillium Therapeutics Inc.
1	University Hospital, Grenoble

# Records	Item Name (19 items)
2	AbbVie
1	Amgen
4	Bristol-Myers Squibb
2	Bristol-Myers Squibb
2	Bristol-Myers Squibb International Corporation
1	Checkpoint Therapeutics, Inc.
1	Exelixis
1	GlaxoSmithKline
1	ImaginAb, Inc.
2	Incyte
1	Incyte Biosciences International Sàl
1	Incyte Corporation
1	MedImmune LLC
2	Merck KGaA
1	Millennium Pharmaceuticals, Inc.
1	NantCell, Inc.
1	NantKwest, Inc.
4	National Cancer Institute
1	Novartis
1	Tarveda Therapeutics
1	Trans-Tasman Radiation Oncology Group
1	Trillium Therapeutics Inc.
1	University Hospital, Grenoble

Technique: Using List Cleanup

How did we create this?

List Cleanup can be used with any type of data. In this recipe, we show list cleanup on company names from a set of clinical trial registry data. You can use the “training.vpxi” file at bizint.com/vptraining to follow the steps in this recipe.

The **training.vpxi** file used in this recipe can be found at:

www.bizint.com/vptraining

VANTAGEPOINT - BIZINT SMART CHARTS EDITION

- Import data:** Open VantagePoint - Smart Charts Edition (VP-SCE) and click **Import from BizInt Smart Charts**. Choose the **training.vpxi** file and click Open.
- When import completes you will see the Summary sheet. Double-click on the company name you wish to cleanup, in this case, the Sponsor field.
- Use Further Processing for initial clean up:** Under the **Refine** tab, select **Further Processing | Read Until First | Paren**. This removes parenthetical information to allow List Cleanup to perform better.
- Start List Cleanup:** Click on **List Cleanup**. You will be prompted to select a matching ruleset to automatically match terms and then go to the Cleanup Confirmation panel to complete the cleanup.
- Choose a matching ruleset:** In this case we’ll choose Organization Names (Pharma) and click **OK**. The Dual View Cleanup Confirmation panel will appear.
- Dual View Cleanup Confirmation:** The left hand panel is sorted by record count and the right is sorted alphabetically. You can filter either list by typing into the filter box. Changes made to either of the two views will update both.
- Display options:** There are different display options available from the drop-down for each of the dual views. The initial display shows **All Items** including what the matching ruleset has combined. Items that are collapsed initially indicate a stronger match, whereas expanded items don’t match as closely and should be reviewed. **Combined Items** shows items that have been combined and **Custom Set of Items** is only used with the Find Close Matches feature. Set the Viewing option to “Combined Items” for the left panel, and to “All items” for the right panel.
- Review combined items:** In this example, the Bristol-Myers Squibb set is expanded in the left panel so you can confirm that these terms were combined correctly. As you review the combined items if you see a group where none of the items should be grouped together, click on the top level item and click the **Delete** key. *This doesn’t delete the individual terms.* To remove a single item from a group, select it and click the Delete key. *Note: The term still exists in the data, but won’t be renamed in the cleaned field.*
- Drag and drop:** Combine groups by dragging one group onto another. Or drag a term from one group to another. In this example, you see that there are two entries for Incyte. Drag and drop the longer name on the shorter name to combine these.
- Filter and combine:** We can use the filter box on either panel to make it easier to find terms to combine. Change the Viewing option on the left panel to “Uncombined Items” and type “Merck” in the filter box. You can then drag “EMD Serono” from the right panel onto “Merck KGaA” in the left panel. *Tip: To check if EMD Serono is now part of this Merck, you can right click on EMD Serono and choose Search Wikipedia or Google to research.*
- Rename a term:** Right-click on a term and choose **Rename Item**. Or, if you prefer to use a different sub term as the group name, right-click on your preferred term and choose **Make This Item the Set Name**. In this example, right click on Incyte Corporation in the right panel and use Rename Item to rename to the shorter “Incyte.”
- Finish Cleanup and Save:** When you have completed cleanup you can click Accept to create the new, cleaned field. *You may save the cleanup as a thesaurus to use with other datasets if you wish.*
- Repeat List Cleanup:** You can continue to clean the field if you wish using **List Cleanup** a second time and choosing a different matching ruleset.
- When you have finished clean up of your data, you can use it in a visualization or you can send the cleaned up fields back to BizInt Smart Charts as new columns using **Export to BizInt Smart Charts** under the **Report** tab.

Technique: Working with Thesauri

VP-SCE offers multiple ways to create and work with thesauri. In many recipes you will notice that when you clean a field with List Cleanup, you have the option to save as a thesaurus. You can also edit and create thesauri directly with the Thesaurus Editor.

Here we show how to edit a Drug Development Thesaurus to change the way that phase terms are grouped and sorted in Excel or VP-SCE visualizations.

The screenshot shows the Thesaurus Editor interface with the following components:

- Field:** A dropdown menu set to "Phase".
- Thesaurus:** A list of terms and their associated regex patterns. The terms include:
 - 0 Discontinued
 - 0 Preclinical
 - 0 Technology
 - 1 Phase 1 (expanded to show patterns like ^Clinical Phase Unknown, ^Clinical\$, etc.)
 - 1 Phase 1/2 (highlighted in blue)
 - 2 Phase 2
 - 3 Phase 3
 - 4 Phase 4
 - 5 Registration
 - 6 Marketed
 - 8 No Development Reported
- Results:** A list of the results after applying the thesaurus, showing:
 - 1 Phase 1
 - 1 Phase 1/2
 - 2 Phase 2
- Display:** Radio buttons for "Entire List" (selected), "Matched Items", and "Un-Matched Items".

Technique: Working with Thesauri

How did we create this?

In this recipe, we use the DrugDevelopmentPhases_ExcelSort.the as a starting point to create a thesaurus that will sort the way we want. We could also create a thesaurus from scratch, or edit a thesaurus created as part of List Cleanup.

VANTAGEPOINT - BIZINT SMART CHARTS EDITION

1. Go to the **Editors** tab and click **Thesaurus Editor** to open the Thesaurus Editor in a new window.
2. Use **File | Open Thesaurus** and select the thesaurus you want to work with. We'll use DrugDevelopmentPhases_ExcelSort.the.
3. Before making changes, save under a new name. From the **File** menu select **Save as**.
4. There are two levels in a VP-SCE thesaurus, the Set and Item. The center pane of the Thesaurus Editor shows the structure of the current thesaurus. When you open a thesaurus only the Set names are shown. These are the terms that will replace anything matching the associated Items when the thesaurus is used. To expand an item and view the Items associated with a Set you can click the plus sign.
5. **Combine Phases:** Drag one Set term under another one to consolidate them. Here we drag "0 Biological Testing" onto "0 Preclinical" so that records with either preclinical or biological testing as the phase will be treated as preclinical in our custom thesaurus.
6. **Change Phase Sorting:** Right-click on the Set name for the phase you want to change and select **Rename Item**. Then change the numbering to reflect your preferred sort order. We right-click on "7 Discontinued" and change the 7 to a 0, moving the Discontinued items from the bottom to the top of the list.
7. You can use outline numbering convention of 1, 1.1, 1.1.1, and so on to have phases sort between the existing numbers. Because of the text descriptions, phases will sort alphanumerically, so standard numbering 10 and above will not sort as expected.
8. **Add a New Phase:** To add a new phase, from the **Thesaurus** menu select **Create New Set** and enter the preferred term that should replace the existing text when the thesaurus is run. For this thesaurus, we use a number followed by a text description for the added phase. We add Phase 1/2 as a separate phase, "1 Phase 1/2".
9. **Check the sort:** Right-click on your new Set name and select **Re-Sort Thesaurus | Ascending**. You can change the number at the front of each phase to change the sort order.
10. **Move the Phase 1/2 terms:** Use ctrl-click to select all the Phase 1/2 Items from the "1 Phase 1" Set and drag and drop them to your new "1 Phase 1/2" Set.
11. You can also add Items to the new phase by right-clicking on the new Set and selecting **Insert New Item**. The drop-down box controls whether you want to match exactly or search within the existing term "contains."
12. Items use regular expressions and can be quite complex, but can also be literal matches. For literal matches begin the Item with a caret (^) and end with a dollar sign (\$).
13. **Check your thesaurus:** Use **File | Load Field** and select a field with matching terms. For this example, choose the Phase field and click the **Apply Thesaurus** button.
14. You'll see the results of applying the thesaurus in the right-hand column. You can also change the radio button on the left-hand side to show only Un-Matched or Matched terms from the field you loaded. This allows you to see what you need to add to your thesaurus.
15. **Save your thesaurus:** Go to **File | Save Thesaurus**. *You can now use your edited thesaurus with recipes in this Cookbook instead of the standard phase thesaurus.*

Technique: Extracting Keywords

You may find yourself wanting to extract a set of keywords from a field. For example, you might want to visualize Endpoint Classifications in search results from Citeline TrialTrove. TrialTrove provides this information as part of the "Study Design" column, in one of two different formats. This Technique shows how to extract a set of keywords from a field in VP-SCE using the My Keywords feature.

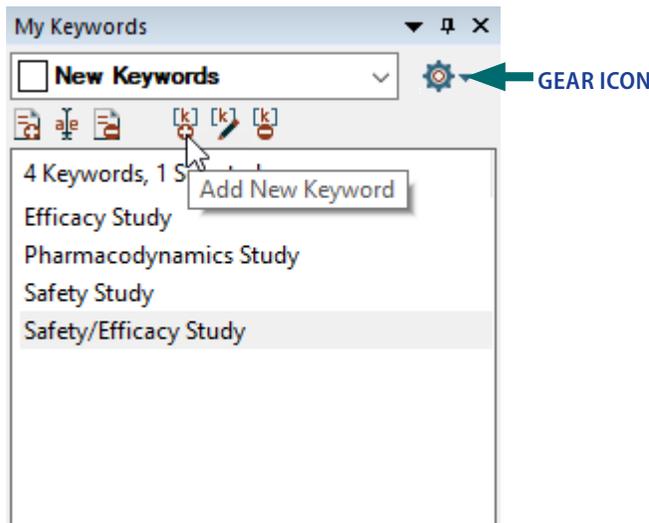


FIGURE 1. THE MY KEYWORDS PANEL

Study Design	Study Design : Study Classifications
This is a double-blind, randomized, active controlled, phase 2, multi-centre, dose ranging, safety study. This was an immunogenicity, single dose study.	Safety Study
Study Type: Interventional Study Design: Allocation: Randomized Control: Active Control Endpoint Classification: Safety/Efficacy Study Intervention Model: Parallel Assignment Masking: Single Blind (Outcomes Assessor) Primary Purpose: Prevention This is a controlled, multicenter, immunogenicity study. This is an observer blinded and tolerability study.	Efficacy Study Safety/Efficacy Study
Study Type: Interventional Study Design: Allocation: Randomized Endpoint Classification: Safety/Efficacy Study Intervention Model: Parallel Assignment Masking: Single Blind (Outcomes Assessor) Primary Purpose: Prevention This is a multicenter, active comparator, immunogenicity study. This is dose response, non-inferiority, tolerability, exploratory study.	Efficacy Study Safety/Efficacy Study

FIGURE 2. KEYWORDS EXTRACTED USING SIMPLE KEYWORDS

Study Design	Study Design : Study Classifications	Study Design : Study Classifications (2)
Study Type: Interventional Study Design: Allocation: Randomized Endpoint Classification: Safety/Efficacy Study Intervention Model: Parallel Assignment Masking: Double-Blind Primary Purpose: Prevention This is an observer blind, multi-center, active controlled, immunogenicity and tolerability study.	Efficacy Study Safety/Efficacy Study	Safety/Efficacy Study

FIGURE 3. KEYWORDS EXTRACTED USING REGULAR EXPRESSIONS

Technique: Extracting Keywords

How did we create this?

SEARCHES:

- For this particular example, we used an export from Citeline TrialTrove. If you would like to follow along, you can import any results from TrialTrove into BizInt Smart Charts Drug Development Suite, make the Study Design column visible, and then export to VP-SCE.

BIZINT SMART CHARTS DRUG DEVELOPMENT SUITE

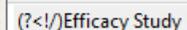
- Import the search results into BizInt Smart Charts to create .cht files.
- Display the “Study Design” column.
- Use **File | Export** to export the chart as VantagePoint - Smart Charts Edition (creates a .vpxi file.)

VANTAGEPOINT - SMART CHARTS EDITION

- If the file does not launch automatically in VantagePoint - Smart Charts Edition (VP-SCE), then import the .vpxi file into VP-SCE using **Import from BizInt Smart Charts** under the **Home** tab.
- Double-click on the Study Design field.
- Make sure the My Keywords panel is visible. Look on the **View** tab and make sure **My Keywords** is selected (in blue). The My Keywords panel is shown in Figure 1 (on the previous page.)
- Click on the **Manage Keywords** blue-red “gear icon” (see Figure 1) and choose **Create New Keyword List**.
- Click the [k]+ **Add New Keyword** button or press Insert on your keyboard to insert a new keyword. Type the phrase you would like to extract from the raw data.
In the example in Figure 1, you can see that we have entered some of the Endpoint Classifications that we see in our data.
- To extract your keywords from a field, bring up the Manage Keywords menu (under the blue-red gear icon), and hold your mouse over **Extract My Keywords From | Study Design**. (Move your mouse over Extract My Keywords From and then move the mouse to the field name that you want to extract from). This will create a new field with only the extracted keywords.
- Export the extracted keywords back to BizInt Smart Charts:** Under the **Report** tab click on **Export to BizInt Smart Charts**. In the Export Wizard, the Study Design : Study Classifications field should already be checked. Click Export.

GOING FURTHER

- Keywords can be case sensitive or not. You can control whether a set of keywords is case sensitive on the Manage Keywords menu (under the blue-red gear icon), and then follow to the **Keyword Options** sub-menu. If you look at the Figure 2 (on the previous page), the first row lists “safety study” in all lower case, but our standard keyword is used in the extracted column.
- My Keywords can do more than simply match phrases. For example, if you look more carefully at the example in Figure 2, you will see that we are extracting both “Safety/Efficacy Study” and “Efficacy Study” from records where the original text contains “Safety/Efficacy Study”. We can use a regular expression to ensure that we only match the exact term “Efficacy Study”.
- To use a regular expression, make sure that there is a check mark next to **Regular Expressions** on the **Keyword Options** sub-menu. Then modify your keyword to include the necessary regular expression code. In this case, we want to only match when the character before “Efficacy” is not a “/”. You can see the regular expression we used in the keyword on the right
- When we Extract My Keywords from the Study Design field again, you see in Figure 3 that we now only extract “Efficacy Study” when that exact text is present, and not when the classification is “Safety/Efficacy Study”.


 (?<!/)Efficacy Study

ABOUT THIS COOKBOOK

This “Cookbook” illustrates how to create some popular visualizations and reports from patent, clinical trials, drug pipeline, and gene sequence data using tools in the BizInt Smart Charts family.

To use these “recipes”, you’ll need the following software:

- BizInt Smart Charts Drug Development Suite 4.0 or higher
- BizInt Smart Charts for Patents 4.0 or higher
- BizInt Smart Charts Reference Rows 2.0 or higher
- VantagePoint - Smart Charts Edition 11.0 or higher

BIZINT SMART CHARTS TRIAL SOFTWARE

You can download BizInt Smart Charts trial software at: <http://www.bizint.com/try>

You can request a trial of VantagePoint - Smart Charts Edition at <http://www.bizint.com/vpsce> or contact us (email: support@bizint.com).

BIZINT SMART CHARTS PRODUCT DOCUMENTATION

You’ll find more information on how to use BizInt Smart Charts software tools at:
<http://www.bizint.com/tips>

Note: These recipes are constantly being tested and revised based on user feedback. We suggest that you check the website to see if there is a more recent version of your recipe. If you find any errors or need help, please contact us at “support@bizint.com.”

APRIL 2019

Download the Software

- **Customer downloads:** <http://www.bizint.com/downloads/>
You will need to enter the userid and password supplied to your company.
- **Trial downloads:** <http://www.bizint.com/try>
We encourage you to evaluate BizInt Smart Charts using your own data!

Keep up with the Latest Releases

- **Latest release information web page:** provides details on all releases.
http://www.bizint.com/product/patents/latest_release.php
http://www.bizint.com/product/drugdevsuite/latest_release.php
- **Subscribe to release announcements** by email or RSS feeds.
<http://www.bizint.com/support/SubscribeReleases.php>

Using the Software

- **Creating Reports from Databases and Hosts web page:** follow the links on the left side of the web page for step by step instructions for each database/host.
<http://www.bizint.com/support/create/>

Go to www.bizint.com/tips for links to these useful documents —

- **Mini Guide:** provides a step-by-step overview of the key features of each BizInt Smart Charts product.
- **Combining and Updating Reports:** a detailed discussion of features and options for creating combined and updated reports.
- **Distributing Reports:** a guide to choosing and using the export formats available with BizInt Smart Charts, including HTML, Word, Acrobat, and Excel.
- **BizInt Smart Charts Reference Rows Mini Guide:** a guide to using the Reference Rows utility for creating single rows combining information from related records.
- **Gene Sequence Reports:** describes the latest features for using BizInt Smart Charts with IP sequence databases.
- **Cookbook of Reports and Visualizations:** www.bizint.com/Cookbook



Contact Us for Help

- **Email us** at support@bizint.com
- **Call us** at +1 (714) 289 1000 (office hours are 8 am - 5 pm Pacific)



Software for
Business Intelligence

BizInt Smart Charts

COOKBOOK OF REPORTS & VISUALIZATIONS

www.bizint.com/Cookbook

The *Cookbook of Reports & Visualizations* illustrates how to create some popular visualizations and reports from drug pipeline, clinical trials, and patent data using tools in the BizInt Smart Charts family.

The Cookbook provides step-by-step recipes showing how you can create these reports and visualizations with your own data.